<u>**AMENDMENTS TO THE CLAIMS**</u>

Claim 1 (Currently Amended)   A video processing apparatus for specifying frames of

content to be start frames of a plurality of viewing segments of the content, when segmenting the

content into the plurality of viewing segments, the video processing apparatus comprising:

a specifying information memory, which is a physical memory storing a plurality of

pieces of specifying information, each piece of specifying information of the plurality of pieces

of specifying information (i) corresponding to a different type of content, and (ii) including:

a first condition identifying a feature of frames of the content to be detected as

candidates for presentation frames, each of the presentation frames for being displayed as a

representative still image of a respective viewing segment of the plurality of viewing segments;

an exclusion condition identifying a feature of frames of the content to be

excluded from being the candidates for the presentation frames;

a second condition identifying a feature of frames of the content to be detected as

candidates for start frames; and

a selection condition identifying a relationship between a presentation frame of

the content and a frame of the content that is to be selected as a start frame;

a content obtaining unit operable to obtain a content;

an information obtaining unit operable to obtain type information identifying the type of

the obtained content;

an extracting unit operable to extract, from the specifying information memory, a piece of

specifying information, of the plurality of pieces of specifying information, that corresponds to

the type of the content identified by the obtained type information; and

a specifying unit operable, in accordance with the extracted piece of specifying information, to (i) specify the presentation frames of the content by detecting, from all frames of the content, frames of the content satisfying the first condition and by subsequently excluding, from the detected frames satisfying the first condition, frames satisfying the exclusion condition, and (ii) specify start frames of the content by detecting, from all frames of the content, frames of the content satisfying the second condition and by subsequently selecting, from the detected frames satisfying the second condition, frames satisfying the relationship identified by the selection condition with respect to the specified presentation frames,

wherein specifying unit includes:

a plurality of detecting subunits, each detecting subunit of the plurality of detecting subunits being operable to detect frames of the content having a different feature;

an excluding subunit operable to exclude, from the detected frames satisfying the first condition, frames satisfying the exclusion condition; and

a selecting subunit operable to select, from the detected frames satisfying the second condition, frames satisfying the relationship identified by the selection condition,

wherein each of the first condition, the exclusion condition, and the second condition is an identifier to be used by one detecting subunit of the plurality of detecting subunits, and

wherein, when operating in accordance with the extracted piece of specifying information corresponding to the type of the content identified by the obtained type information, the specifying unit (i) detects, from all frames of the content, a plurality of large-caption start frames, each large-caption start frame, of the plurality of large-caption start frames, being a first frame of a series of frames of the content during which a caption of a size larger than a threshold continuously appears in a predetermined region, and (ii) specifies, as a presentation frame of the

3

content, each large-caption start frame, of the plurality of large-caption start frames, remaining after removing, from the plurality of large-caption start frames, small-caption frames having a caption of a size smaller than a threshold appearing in a region other than the predetermined region.

Claim 2 (Cancelled)

Claim 3 (Previously Presented)   The video processing apparatus of Claim 1, further comprising an index storage unit operable to store, in correspondence with the content, a respective display time of each of the specified start frames of the content and each of the specified presentation frames of the content.

Claims 4-7 (Cancelled)

Claim 8 (Currently Amended)   The video processing apparatus of Claim 1 ~~Claim 7~~, wherein, when operating in accordance with the extracted piece of specifying information corresponding to the type of the content identified by the obtained type information, the specifying unit (i) detects, from all frames present in the content, a plurality of transition frames, each transition frame, of the plurality of transition frames, being (a) a first frame of a series of frames of the content of similar images, (b) silent frames of the content having audio data below a predetermined volume level, (c) music-start frames of the content having a first frame of a series of frames of the content having audio data representing a piece of music data, or (d) speech-start frames having a first frame of a series of frames of the content having audio data

4

representing a speech of a specific speaker and (ii) specifies as a start frame of the content, for each presentation frame of the content, a frame of the content closest to the presentation frame among all of the detected frames preceding the presentation frame.

Claims 9-15 (Cancelled)

Claim 16 (Previously Presented)   The video processing apparatus of Claim 1, further comprising a playback unit operable to play back the content starting from a start frame of the content specified by the specifying unit.

Claim 17 (Previously Presented)   The video processing apparatus of Claim 16,
        wherein the video processing apparatus further comprises:
            an index storing unit operable to store pairs of display times of each start frame and presentation frame specified for a respective viewing segment, of the plurality of viewing segments, by the specifying unit;
            a display unit operable to display a presentation frame specified for each viewing segment, of the plurality of viewing segments, by the specifying unit; and
            a user-selection unit operable to select, in accordance with a user selection, at least one of the displayed presentation frames, and
        wherein the playback unit plays back the content starting from a start frame of a viewing segment, of the plurality of viewing segments, to which the user-selected presentation frame belongs.

Claim 18 (Previously Presented)   The video processing apparatus of Claim 17, wherein the display unit displays the presentation frames specified for the plurality of viewing segments by generating a thumbnail image of each presentation frame of the presentation frames of the content and displaying the generated thumbnail images in list form.

Claim 19 (Previously Presented)   The video processing apparatus of Claim 17,

wherein the user-selection unit stores the at least one of the selected presentation frames as a reference image into the specifying information memory, and

wherein the specifying unit specifies the presentation frames by detecting frames of the content similar to the reference image with respect to a location of a region in which a caption appears.

Claim 20 (Previously Presented)   The video processing apparatus of Claim 1,

wherein the video processing apparatus includes a recording unit operable to obtain a content and type information of the content, and to record the content to a recording medium in correspondence with the type information,

wherein after the recording unit records the type information and at least a part of the content, the content obtaining unit sequentially obtains the part of the content from the recording medium, and

wherein the specifying unit sequentially specifies a start frame present in the part of the content obtained by the content obtaining unit.

Claim 21 (Previously Presented)   The video processing apparatus of Claim 1,

wherein the video processing apparatus comprises a recording unit operable to obtain a content and type information of the content, encode the content, and record the encoded content in correspondence with the type information,

wherein, after the recording unit records the type information and encodes at least a part of the content, the content obtaining unit sequentially obtains the encoded part of the content, and

wherein the specifying unit obtains an analysis of the encoded part of the content, and sequentially specifies a start frame present in the encoded part of the content using the obtained analysis.

Claim 22 (Previously Presented)   The video processing apparatus of Claim 1, further comprising an updating unit operable to obtain a new version of a piece of specifying information, of the plurality of pieces of specifying information, that corresponds to a specific type of content, and operable to record the new version of the piece of specifying information to the specifying information memory.

Claim 23 (Previously Presented)   The video processing apparatus of Claim 22,

wherein the updating unit obtains the new version of the piece of specifying information when the video processing apparatus is connected, via a communication network, to a provider apparatus for providing specifying information, and judges that the new version of the piece of specifying information is available, and

wherein the new version of the piece of specifying information is recorded to the specifying information memory by updating, to the new version of the piece of specifying

information, a stored piece of specifying information, of the plurality of pieces of specifying information, that corresponds to the specific type.

Claim 24 (Previously Presented)   The video processing apparatus of Claim 23, wherein a judgment as to whether the new version of the piece of specifying information is available is made each time the specifying unit processes the specific type of content.

Claim 25 (Currently Amended)   An integrated circuit for use in a video processing apparatus that specifies frames of content to be start frames of a plurality of viewing segments of the content, when segmenting the content into the plurality of viewing segments, the video processing apparatus having a specifying information memory storing a plurality of pieces of specifying information, each piece of specifying information of the plurality of pieces of specifying information corresponding to a different type of content and including (i) a first condition identifying a feature of frames of the content to be detected as candidates for presentation frames, each of the presentation frames for being displayed as a representative still image of a respective viewing segment of the plurality of viewing segments, (ii) an exclusion condition identifying a feature of frames of the content to be excluded from being the candidates for the presentation frames, (iii) a second condition identifying a feature of frames of the content to be detected as candidates for start frames, and (iv) a selection condition identifying a relationship between a presentation frame of the content and a frame of the content that is to be selected as a start frame, the integrated circuit comprising:

a content obtaining module operable to obtain a content;

an information obtaining module operable to obtain type information identifying the type of the obtained content;

an extracting module operable to extract, from the specifying information memory, a piece of specifying information, of the plurality of pieces of specifying information, that corresponds to the type of the content identified by the obtained type information; and

a specifying module operable, in accordance with the extracted piece of specifying information, to (i) specify the presentation frames of the content by detecting, from all frames of the content, frames of the content satisfying the first condition and by subsequently excluding, from the detected frames satisfying the first condition, frames satisfying the exclusion condition, and (ii) specify start frames of the content by detecting, from all frames of the content, frames of the content satisfying the second condition and by subsequently selecting, from the detected frames satisfying the second condition, frames satisfying the relationship identified by the selection condition with respect to the specified presentation frames,

wherein specifying module includes:

a plurality of detecting submodules, each detecting submodule of the plurality of detecting submodules being operable to detect frames of the content having a different feature;

an excluding submodule operable to exclude, from the detected frames satisfying the first condition, frames satisfying the exclusion condition; and

a selecting submodule operable to select, from the detected frames satisfying the second condition, frames satisfying the relationship identified by the selection condition,

wherein each of the first condition, the exclusion condition, and the second condition is an identifier to be used by one detecting submodule of the plurality of detecting submodules, and

wherein, when operating in accordance with the extracted piece of specifying information corresponding to the type of the content identified by the obtained type information, the specifying module (i) detects, from all frames of the content, a plurality of large-caption start frames, each large-caption start frame, of the plurality of large-caption start frames, being a first frame of a series of frames of the content during which a caption of a size larger than a threshold continuously appears in a predetermined region, and (ii) specifies, as a presentation frame of the content, each large-caption start frame, of the plurality of large-caption start frames, remaining after removing, from the plurality of large-caption start frames, small-caption frames having a caption of a size smaller than a threshold appearing in a region other than the predetermined region.

Claim 26 (Currently Amended)   A video processing method for use by a video processing apparatus that specifies frames of content to be start frames of a plurality of viewing segments of the content, when segmenting the content into the plurality of viewing segments, the video processing apparatus having a specifying information memory storing a plurality of pieces of specifying information, each piece of specifying information of the plurality of pieces of specifying information corresponding to a different type of content and including (i) a first condition identifying a feature of frames of the content to be detected as candidates for presentation frames, each of the presentation frames for being displayed as a representative still image of a respective viewing segment of the plurality of viewing segments, (ii) an exclusion condition identifying a feature of frames of the content to be excluded from being the candidates for the presentation frames, (iii) a second condition identifying a feature of frames of the content to be detected as candidates for start frames, and (iv) a selection condition identifying a

relationship between a presentation frame of the content and a frame of the content that is to be selected as a start frame, the video processing method comprising:

obtaining a content;

obtaining type information identifying the type of the obtained content;

extracting, from the specifying information memory, a piece of specifying information, of the plurality of pieces of specifying information, that corresponds to the type of the content identified by the obtained type information;

specifying, in accordance with the extracted piece of specifying information, presentation frames of the content by detecting, from all frames of the content, frames of the content satisfying the first condition and by subsequently excluding, from the detected frames satisfying the first condition, frames satisfying the exclusion condition; and

specifying, in accordance with the extracted piece of specifying information, start frames of the content by detecting, from all frames of the content, frames of the content satisfying the second condition and by subsequently selecting, from the detected frames satisfying the second condition, frames satisfying the relationship identified by the selection condition with respect to the specified presentation frames,

detecting, via each of a plurality of detecting subunits, frames of the content having a different feature;

excluding, via an excluding subunit and from the detected frames satisfying the first condition, frames satisfying the exclusion condition; and

selecting, via a selecting subunit and from the detected frames satisfying the second condition, frames satisfying the relationship identified by the selection condition,

wherein each of the first condition, the exclusion condition, and the second condition is an identifier to be used by one detecting subunit, of the plurality of detecting subunits, when performing the detecting, and

wherein, when operating in accordance with the extracted piece of specifying information corresponding to the type of the content identified by the obtained type information, the video processing method (i) detects, from all frames of the content, a plurality of large-caption start frames, each large-caption start frame, of the plurality of large-caption start frames, being a first frame of a series of frames of the content during which a caption of a size larger than a threshold continuously appears in a predetermined region, and (ii) specifies, as a presentation frame of the content, each large-caption start frame, of the plurality of large-caption start frames, remaining after removing, from the plurality of large-caption start frames, small-caption frames having a caption of a size smaller than a threshold appearing in a region other than the predetermined region.

Claim 27 (Currently Amended)　　A non-transitory computer-readable recording medium have a video processing program recorded thereon, the video processing program for causing a device to specify frames of content to be start frames of a plurality of viewing segments of the content, when segmenting the content into the plurality of viewing segments, the device having a specifying information memory storing a plurality of pieces of specifying information, each piece of specifying information of the plurality of pieces of specifying information corresponding to a different type of content and including (i) a first condition identifying a feature of frames of the content to be detected as candidates for presentation frames, each of the presentation frames for being displayed as a representative still image of a respective viewing

segment of the plurality of viewing segments, (ii) an exclusion condition identifying a feature of frames of the content to be excluded from being the candidates for the presentation frames, (iii) a second condition identifying a feature of frames of the content to be detected as candidates for start frames, and (iv) a selection condition identifying a relationship between a presentation frame of the content and a frame of the content that is to be selected as a start frame, the video processing program causing a computer to execute a method comprising:

obtaining a content;

obtaining type information identifying the type of the obtained content;

extracting, from the specifying information memory, a piece of specifying information, of the plurality of pieces of specifying information, that corresponds to the type of the content identified by the obtained type information;

specifying, in accordance with the extracted piece of specifying information, presentation frames of the content by detecting, from all frames of the content, frames of the content satisfying the first condition and by subsequently excluding, from the detected frames satisfying the first condition, frames satisfying the exclusion condition; and

specifying, in accordance with the extracted piece of specifying information, start frames of the content by detecting, from all frames of the content, frames of the content satisfying the second condition and by subsequently selecting, from the detected frames satisfying the second condition, frames satisfying the relationship identified by the selection condition with respect to the specified presentation frames,

detecting, via each of a plurality of detecting subunits, frames of the content having a different feature;

excluding, via an excluding subunit and from the detected frames satisfying the first condition, frames satisfying the exclusion condition; and

selecting, via a selecting subunit and from the detected frames satisfying the second condition, frames satisfying the relationship identified by the selection condition,

wherein each of the first condition, the exclusion condition, and the second condition is an identifier to be used by one detecting subunit, of the plurality of detecting subunits, when performing the detecting, and

wherein, when operating in accordance with the extracted piece of specifying information corresponding to the type of the content identified by the obtained type information, the method executed by the computer (i) detects, from all frames of the content, a plurality of large-caption start frames, each large-caption start frame, of the plurality of large-caption start frames, being a first frame of a series of frames of the content during which a caption of a size larger than a threshold continuously appears in a predetermined region, and (ii) specifies, as a presentation frame of the content, each large-caption start frame, of the plurality of large-caption start frames, remaining after removing, from the plurality of large-caption start frames, small-caption frames having a caption of a size smaller than a threshold appearing in a region other than the predetermined region.